

STATISTICAL METHODOLOGIES TO INTEGRATE EXPERIMENTAL AND COMPUTATIONAL RESEARCH *

P. A. Parker
National Aeronautics and Space Administration
Langley Research Center
Hampton, VA

R. T. Johnson and D. C. Montgomery
Arizona State University
Industrial Engineering
Tempe, AZ

ABSTRACT

Development of advanced algorithms for simulating engine flow paths requires the integration of fundamental experiments with the validation of enhanced mathematical models. In this paper, we provide an overview of statistical methods to strategically and efficiently conduct experiments and computational model refinement. Moreover, the integration of experimental and computational research efforts is emphasized. With a statistical engineering perspective, scientific and engineering expertise is combined with statistical sciences to gain deeper insights into experimental phenomenon and code development performance; supporting the overall research objectives. The particular statistical methods discussed are design of experiments, response surface methodology, and uncertainty analysis and planning. Their application is illustrated with a coaxial free jet experiment and a turbulence model refinement investigation. Our goal is to provide an overview, focusing on concepts rather than practice, to demonstrate the benefits of using statistical methods in research and development, thereby encouraging their broader and more systematic application.

INTRODUCTION

Research efforts to improve the ability to test hypersonic vehicles seek a better understanding of the test media effects from ground test facilities. These efforts include computer modeling, experimentation, and diagnostics. More specifically, the primary objectives are (1) obtain a better understanding of facility effects through computer modeling, experimentation, and diagnostics, (2) develop enhanced codes with increased capability to model turbulence, turbulent mixing, and kinetics, (3) improve diagnostics for increased fidelity experimental measurements, (4) conduct fundamental experiments to be used in model development and code validation¹. In each of these research objectives, there is an opportunity to apply powerful statistical methods to enhance current practices and provide a systematic and defensible framework to plan and conduct an investigation, identify sources of variability. As a result, the methods support scientific conclusions with a quantified and defensible level of confidence. While the use of statistical techniques is routine in post-experimental data analysis, we propose a higher-level view that integrates the engineering and scientific expertise with statistical sciences to strategically and efficiently plan and conduct the investigation to meet the research objectives. A unifying framework of this nature is particularly useful in a distributed research environment to integrate multiple code development and experimental activities. In addition, statistical engineering helps to establish quantitative interfaces between the focused research efforts to support the overall objectives.

Research and development processes and systems can be represented by a simplified diagram shown in Figure 1. The box in the center of the diagram contains the system under investigation, which could be a physical experiment, a computational investigation, or the integration of experimental and

* Approved for public release; distribution is unlimited.

computational results. The system is acted upon by factors, x 's, and produces responses, y 's, and the functional form, $f(x)$, mathematically describes the factor-response relationship. In this research, there is considerable physics-based knowledge about the system under investigation, and therefore we seek to understand where our predictions of physical phenomenon deviate from experimental data.

To gain a better understanding, we conduct an experimental or computational investigation to interrogate the system, generating observations that are tested against our current predictive capability. Therefore, the purpose of the investigation is to identify the magnitude of the uncertainty in our postulated functional relationship and gain knowledge about the influence of the factors on that uncertainty. An interrogation of the system, often referred to as collecting a data point, can be costly in either time and/or expense. Therefore with limited resources, it is critical to consider the amount of information gained, or benefit, from each data point. Experimental efficiency is defined as the amount of information gained per data point. By forming precise objectives and research questions, usually in the form of hypothesis tests, statistical methods can quantitatively assess efficiency and specify adequate data volume. Intuitively, high efficiency can be gained by reducing the cost of experimentation, however in this paper we restrict our attention to increasing efficiency by strategically specifying factor combinations which are information-rich, assuming that the cost per data point is fixed. Applying statistical thinking at the system level provides an efficient means to interrogate, quantify, isolate, and model the uncertainty in the knowledge about the physics-based functional relationships.

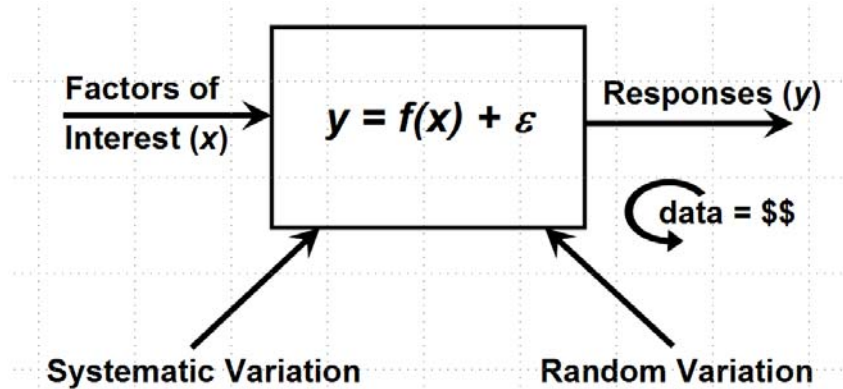


Figure 1: Systems View

From Figure 1, we categorize uncertainty into systematic and random variation. In an experimental investigation, systematic sources include facility warming and cooling trends over the testing time, uncharacterized geometric and boundary conditions in the flow field, and biases in the diagnostic measurement systems. All of these sources produce a predictable influence on the system, however we are either unable to fully correct for their impact or they are not of primary interest to the investigation. In contrast, random variation can not be predicted and includes variability in setting the experimental flow-field conditions and the precision of diagnostic measurements. While some of these sources of variation are known and others are unknown, they are distinguished from factors in that we do not, or cannot, model their influence on the system. To draw research-oriented conclusions from the experimental data about the deviations from our predictions, we must discern, in the presences of uncertainty, the discovery of new knowledge about the physical phenomenon and experimental variation. A statistical viewpoint recognizes and plans for the presences of various sources of variability, and supports an objective framework that allows researchers to make rigorous decisions and inferences in the presence of uncertainty, thereby validating their scientific conclusions. In the following sections, we illustrate the connection between the research objectives and applicable statistical methods and illustrate their application.

RESULTS AND DISCUSSION

OVERVIEW OF STATISTICAL METHODS

There are three statistical methods that are directly applicable to the research objectives, namely (1) Design of Experiments, (2) Response Surface Methodology, and (3) Uncertainty Analysis. These methods are widely utilized in industry for product and process optimization, with a particular emphasis on understanding variability and accelerating development time²⁻⁵. In addition, these methods have been demonstrated to be beneficial in various aspects of previous hypersonic propulsion research⁶⁻⁹. Therefore, our goal is not introduce new statistical methods or to initiate their application to hypersonic research, rather it is to provide an overview of their broad applicability and encourage more systematic utilization.

In the broadest sense, a designed experiment is a purposeful control of the inputs (factors) in such a way as to deduce their relationship, if any, with the outputs (responses). Statistical design of experiments is the process of planning an experiment so that appropriate data are collected to answer research questions with valid and objective conclusions. Design of experiments incorporates prior engineering and scientific knowledge to plan and conduct an experiment with a careful attention to experimental efficiency. With an emphasis on extensive planning before execution, a well-designed experiment can help to ensure that research questions are answered with a specified level of confidence. This emphasis on quantitatively assessing the experiment's performance in advance of execution is a distinguishing aspect of a statistically designed experiment. In general, an experiment design specifies the levels of the input factors and employs tactical execution techniques to concurrently collect data and isolate sources of variability. The primary execution techniques are randomization data point collection, replication of experimental conditions, and blocking into relatively homogenous experimental periods. Using these methods enables insightful analyses to identify sources of variability and partition them into nuisance components and those that are of research interest. The concepts and tools used to statistically design experiments are applicable to physical experiments and computational investigations.

Response surface methodology (RSM) is a collection of statistical modeling techniques for studying, characterizing, improving, and optimizing processes. For example, we may seek to mathematically model the response parameters in an experimentally measured flow-field to infer measurements at intermediate locations in the design space, where actual data were not collected. A distinction between RSM and typical curve-fitting is a desire to build parsimonious models that adequately capture the functional relationship within the experimental error, thereby defending against over-fitting which fails to recognize random sources of variability and ultimately introduces more uncertainty into data-driven conclusions. As another example, RSM is used to compare the response surfaces derived from experimental data to the computational response surfaces from simulation codes. In this case, mathematical models are built of the experimental-to-computational agreement as a function of model tuning constants to improve the correlation to experimental results over multiple configurations and flow conditions.

Uncertainty analysis is employed in the planning phase to estimate the data volume required to meet the research objectives. The result of uncertainty analysis is a specification of a sufficient sampling strategy. While each measurement from a diagnostic system has an associated precision limited by the instrument, averaging of measurements increases the precision of the flow-field parameters to a specified precision. Conceptually, this is straightforward; however in practice the uncertainty analysis can become quite complex due to the presence of multiple sources of variability and the requirement to estimate higher-order distributional parameters, such as variances and covariances. The process begins with a careful partitioning of the known sources of variability and requires the balancing of competing criteria due to limited experimental resources. Statistical uncertainty analysis and planning provides valuable insights before the experiment is conducted that are helpful in the analysis and it guides the decisions on the number of measurements, or observations, required to obtain the specified fidelity of the flow-field parameters.

Synergistically combining these statistical methods with scientific and engineering knowledge supports a better understanding of facility effects through computer modeling, experimentation, and diagnostics. While the approach to design a simple experiment is often straightforward, the dimensionality and complexity of the sources of variability of the applications we consider is a compelling reason to employ these rigorous methods.

COAXIAL FREE JET EXPERIMENT EXAMPLE

To illustrate the application of statistical methods, consider an experiment to characterize the flow-field parameters as a function of the location within a coaxial free jet to study turbulent mixing¹⁰. Design of experiments and uncertainty analysis planning were applied to create simultaneous response surfaces of twenty-seven flow parameters as a function of the axial and radial distance within the coaxial free jet. A drawing and infrared photo illustrate the of the coaxial free jet configuration shown in Figure 2. A list of the factors and response variables is provided in Table 1. In addition to the variability of the responses, the pair-wise covariances are also of interest. A combined coherent anti-Stokes Raman spectroscopy¹¹ and interferometric Rayleigh scattering system is used as the diagnostic measurement technique¹¹.

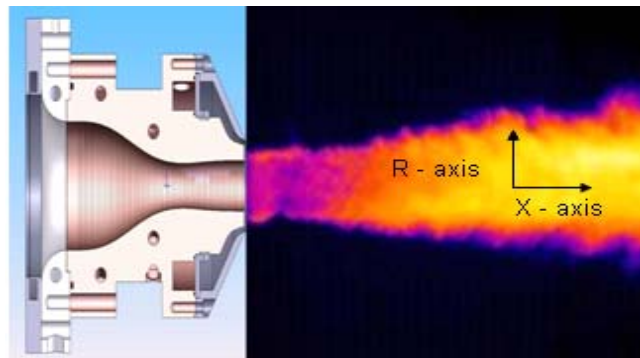


Figure 2: Infrared image of the coaxial free jet experiment.

Table 1: Factors and responses in the coaxial free jet experiment.

Factors	Responses
X – location R – location	Temperature u – axis velocity v – axis velocity N ₂ (%) H ₂ (%) O ₂ (%) Var (T), Var (u), Var (v), Var (N ₂), Var (H ₂), Var (O ₂)

Several design challenges are: What type of experimental design, which specifies the locations in the flow-field to collect data, should be used?; How will the number of unique design points be determined?; How many measurements should be taken at a design point in order to account for measurement noise and estimate turbulence? To meet the experimental objectives and answer the design challenge questions posed, a classical design approach was used. In this example, a face-centered central composite design (FCCD) was chosen as the base experiment, which consists of nine design points on the corners, edge-centers, and center of a square. This design is commonly used for

response surface modeling⁵ and provides the ability to fit second-order response models. Due to the complex nature of the experiment it is anticipated that a global model over the entire domain would be greater than second order, however over small regions it could be approximated with a second order model. Therefore, multiple nested FCCD designs were placed over the entire domain, guided by the predicted characteristics of the flow-field. The nesting of the FCCD designs supports piecewise modeling of second-order models in sub-sections of the design space along with higher order polynomial models across larger regions. The near uniformity of the design also allowed for non-parametric model fitting. Design points were also added in predicted areas of highest variability in the turbulent flow so that high fidelity models could be constructed to capture steep gradient along specific radial traces. The overlaying of these multiple criteria resulted in the design shown in Figure 3.

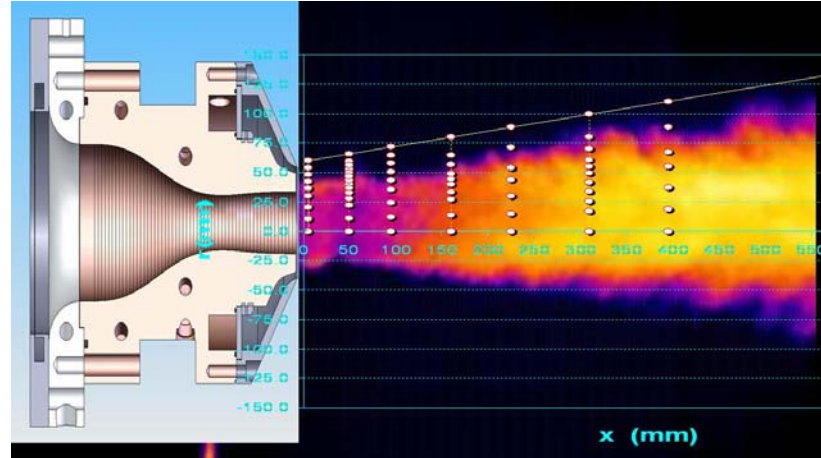


Figure 3: Design point locations for coaxial free jet experiment

A statistical design approach not only specifies the locations to obtain measurements, but also the experimental protocol in which they are collected. The three basic principles design of experiment execution are randomization, replication, and blocking⁴. Randomization of the run ordering supports statistical independence of observations, an assumption in typical regression analysis, and defends against systematic trending sources of variation. Replication refers to the repeat of experimental design points after resetting the flow-field conditions; not simply repeating measurements at a particular location. Replication provides pure-error estimates of experimental error, which includes the variability in the measurement system and the experimental conditions. Blocking specifies a collection of design points to be executed under relatively homogenous conditions. In this experiment, the block was defined by the amount of time of a single run of the coaxial free jet apparatus. Due to this experimental limitation, it was important to divide the experiments in a strategic manner that allows for the estimation of block effects, namely those due to run-to-run or day-to-day variations in the experimental error over time. To specify the order of execution, the design points were completely randomized by sub-region in the flow-field. Then, blocks were defined in a way that ensured repeated design points (replicates) were contained in a sampling of blocks. While more sophisticated blocking strategies are available, this relatively simple approach accommodated this sources of variability in this example apparatus.

At this point, we have covered the choice of design locations, replication of the design locations, randomization of run ordering, and blocking strategies. We now consider the number of measurements to be taken at each location, referred to as the sub-sampling strategy. In this experiment, the estimation of variability in the flow-field parameters is of primary interest, and a rigorous uncertainty analysis was used to specify the number of sub-samples. Conceptually, it is clear that we require a large number of observations to estimate the variance of a distribution as compared to a mean, since it is a higher-order parameter. For each set of sub-samples taken at a single location, we obtain an estimate of the variability (turbulence) in the flow-field. Furthermore, we desire a model of how the turbulence changes over the domain of the flow-field. It is instructive to partition the various sources of variability in the

mathematical relationship estimated from the data. For example, we can express the standard deviation of the temperature as a function of the x and r locations in the flow-field as

$$\sigma_T = f(x, r) + \varepsilon ,$$

where $f(x, r)$ is the functional relationship describing how the temperature variability changes over the flow-field, and ε is residual error, or the unexplained deviations of the model from experimental data. Consider the sources that contribute to the unexplained variance as

$$\text{var}(\varepsilon) = \sigma_\varepsilon^2 = \sigma_{\text{pure-error}}^2 + \sigma_{\text{lack-of-fit}}^2 ,$$

where, the pure-error component describes the experimental variability, which is model independent, and the lack-of-fit component is due to systematic deviations from the functional form from of the true underlying response surface. Pure-error represents the variability in the parameters when the flow-field parameters are set to the same conditions and provides an objective guide to prevent over-fitting in the model. Conceptually, if lack-of-fit is large relative the pure-error, then a more complex model could be considered. Alternatively, increasing the order or complexity of the model would not be justified and would result in the undesirable influence of random noise. The pure-error component can be further partitioned into

$$\sigma_{\text{pure-error}}^2 = \sigma_{\text{experimental settings}}^2 + \sigma_{\text{measurement precision}}^2 ,$$

which includes the variability in setting flow-field and the measurements precision. The precision in estimating the standard deviation of temperature depends on the number of replications of design locations and the number of sub-samples chosen in the experimental design. Using uncertainty analysis, we can estimate the required data volume (number of measurements) at a location to achieve a specified precision in estimating the standard deviation of the parameter based on the pre-experimental knowledge of the experimental variance components.

While the data in this example is not available, we briefly discuss the planned response surface modeling approach. To increase our understanding of the entire flow-field space, it is important to produce a mathematical model that relates the factors (x, r) to the parameters (responses), and most importantly identifies the areas where our predictions deviate from the experimental data. Figure 4 illustrates a predicted response surface based on simulation and highlights the steep gradients that are anticipated in the flow-field. As previously mentioned, several modeling strategies are considered and supported by the experimental design to capture these gradients. The two primary methods considered are (1) a parametric polynomial model based on a Taylor series expansion and (2) a non-parametric Gaussian process model. Each method has strengths and weaknesses. For example, a parametric model provides a convenient and easily interpretable functional relationship, however for response surfaces with steep gradients it requires many high-order terms. Alternatively, a Gaussian process model is essentially an interpolating function that can provide an excellent fit to the experimental response surface, however over-fitting is a concern. Since the experimental designs considers both modeling strategies, in the analysis of the experimental data a comparison of these methods can be performed.

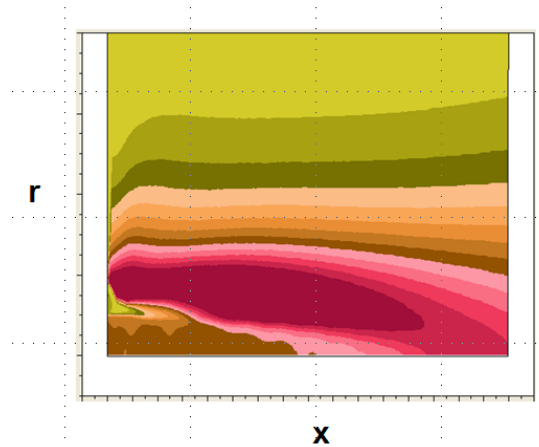


Figure 4: Predicted temperature response as a function of x and r locations

COMPUTATIONAL MODEL REFINEMENT EXAMPLE

Improved numerical simulation of hypersonic engine performance relies on the development of enhanced codes with an increased capability to model turbulence, turbulent mixing, and kinetics. The validation phase of these numerical models requires the comparison of simulated results to experimental data obtained from multiple experimental cases. One aspect of validation, referred to as model refinement, involves the selection of model tuning parameters (coefficients) to achieve the *best* agreement with the experimental data.¹²

The model refinement phase is essentially an experiment in which we set specific values of the model tuning coefficients, execute simulation cases, quantify the correlation to experimental results, and seek values of the coefficients that improve the correlation, or agreement, with actual experimental measurements. In the RSM context, the model tuning parameters are the factors and the simulation-to-experimental correlation are the responses that we seek to improve. Applying RSM offers a systematic approach to perform model refinement that emphasizes the use of minimal computational resources and features an analysis approach to gain deeper insights into the underlying physics.

The procedure involves the selection of the particular coefficients to vary, their ranges, and developing an experimental design that specifies the levels and combinations of factors to be run through the simulation model. For each combination of factors, known as an experimental run, measures of simulation-to-experimental agreement (responses) are obtained. For a computational experiment, a class of experimental designs known as space-filling are being considered. These designs provide a relatively uniform distribution of information throughout the design space, which is particularly valuable when parametric assumptions about the response surface are weak.

From these data, validation models of the relationship between the factors and the responses are estimated. These validation models describe a multidimensional validation surface of the difference between the simulated and experimental results as a function of the model tuning coefficients. Our goal is to identify the regions of the validation surface (combination of factors) that represent agreement with the experimental results.

Since the functional form of the validation surface is unknown, an iterative process is performed to assess the adequacy of an estimated validation model and augment the experimental design as required, thereby enabling the estimation of a higher fidelity model. Once adequate models are found, they are combined to perform multiple response optimization, thereby estimating the values of the model tuning parameters to achieve the best correlation to the experimental data for all of the response quantities of interest. Note that we use the term *best* to describe a trade-off among multiple competing

criteria in correlating different components of the simulation model exercised by different flow-field configurations.

In the final step, the values of the model tuning parameters determined by the optimization are run through the simulation code to confirm that the predicted quality of correlation with the experimental results is obtained. This confirmation phase provides confidence in the estimated validation models and the optimization results.

In summary, the proposed RSM approach to model refinement is expected to offer a general, structured approach to obtain values for the model tuning parameters. In addition, due to the structured nature of the approach, the selection of the tuning parameters will be reproducible by other researchers. A particular strength of the approach is its straightforward extension to higher-dimensional factor spaces and its ability to incorporate multiple experimental cases, thereby providing a set of parameters that are adequate over a range of experimental conditions.

SUMMARY AND CONCLUSIONS

In this paper, a broad overview of the applicability of statistical methodologies to engine flow path research, particularly emphasizing the integration of experimental and computational efforts, is provided. More specifically, we have discussed an approach to combine scientific and engineering expertise with statistical sciences to increase the efficiency in gaining new knowledge that advances the understanding of the physical phenomenon. Partitioning the sources of variability and deviation of simulation predictions from experimental results enables new and deeper insights to combustion model development. The techniques of design of experiments, response surface methodology, and uncertainty analysis form a systematic framework to plan, execute, and analyze experimental investigations. Moreover, they focus on the strategic allocation of resources to interrogate systems and make valid conclusions in the presence of uncertainty. Through the examples presented, the general applicability of these powerful statistical tools is illustrated. We encourage more systematic and strategic utilization of these methods to enhance research efforts, particularly in high-dimensional design space and under the constraint of limited resources.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support of Rob Baurle, Paul Danehy, Phil Drummond, and Rick Gaffney from NASA Langley Research Center and Andrew Cutler from George Washington University, for their interest, support, and application of statistical methods in their research efforts.

REFERENCES

1. Drummond, P., Danehy, P., Bivolaru, D., Gaffney, R., Parker, P., Chelliah, H., Cutler, A., Givi, P., and Hassan, H. (2006), "Predicting the Effects of Test Media in Ground-Based Propulsion Testing," *Annual ITEA Technology Review*, Technology for High Speed and Hypersonic T & E.
2. Box, G.E.P., Hunter, W.G., Hunter, J.S. (1978), *Statistics for Experimenters*, John Wiley & Sons
3. Box, G.E.P. and Draper, N.R. (1987), *Empirical Model-Building and Response Surfaces*, John Wiley & Sons
4. Montgomery, D.C. (2004), *Design and Analysis of Experiments* (6th ed.), John Wiley & Sons.
5. Myers, R.H. and Montgomery, D.C. (2002), *Response Surface Methodology*, (2nd Ed.) John Wiley & Sons.
6. Cutler, A. D., Danehy, P. M., Byrne, P.O., Rodriguez, C.G., and Drummond, J. P., (2004) "Supersonic Combustion Experiments for CFD Model Development and Validation," AIAA 2004-0266.

7. Cutler, A. D., Diskin, G. S., Danehy, P. M., and Drummond, J. P., (2002) "Fundamental Mixing and Combustion Experiments for Propelled Hypersonic Flight," AIAA 2002-3879.
8. Danehy, P.M., DeLoach, R. Cutler, A.D. (2002) "Application of Modern Design of Experiments to CARS Thermometry in a Supersonic Combustor," AIAA-2002-2914.
9. McClinton, C.R., Ferlemann, S.M., Rock, K.E., Ferlemann, P.G. (2002) "The Role of Formal Experiment Design in Hypersonic Flight System Technology Development," AIAA-2002-0543.
10. Cutler, A.D., Magnotti, G., Baurle, R., Bivolaru, D., Tedder, S.A., Danehy, P.M., Weigl, M.C., Beyrau, F., and Seeger, T. (2007) Development of Supersonic Combustion Experiments for CFD Modeling," AIAA-2007-978.
11. Tedder, S., Bivolaru, D., Danehy, P., Weigl, M., Beyrau, F., Seeger, T., and Cutler, A. (2007) "Characterization of a Combined CARS and Interferometric Rayleigh Scattering System," AIAA-2007-871.
12. Xiao, X., Edwards, J.R., and Hassan, H.A. (2006) "Variable Turbulent Schmidt-Number Formulation for Scramjet Applications," AIAA Journal, 44, pp. 593-599.